What observations do we need for convective-scale data assimilation?

Martin Weissmann, S. Geiß, T. Necker and M. Sommer
Hans-Ertel Centre for Weather research

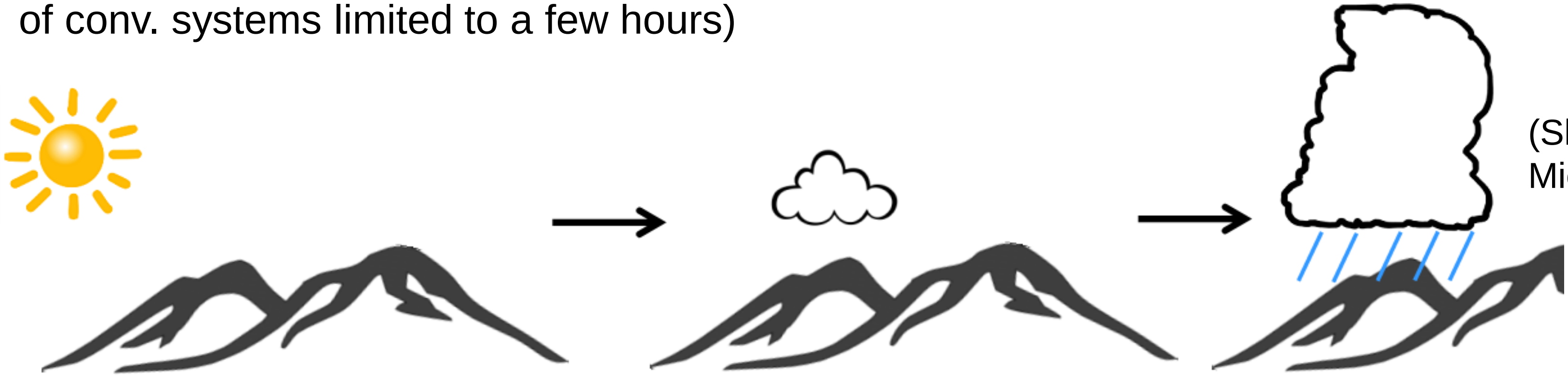
Research in cooperation with the group of R. Potthast at DWD

Regional convective-scale NWP

- Focus on regional extreme events related to convection
- Convection-permitting resolution (km-scale)
- Rapid evolution and low predictability (life-time of conv. systems limited to a few hours)

Suitable observations

- Observations related to convection (environment, clouds, precipitation)
- High spatial resolution (1 – 10 km)
- High temporal resolution (5 – 15 min)



(Sketch from Michael Keller)

Environmental conditions (T, p, wind, q)

Cloud observations

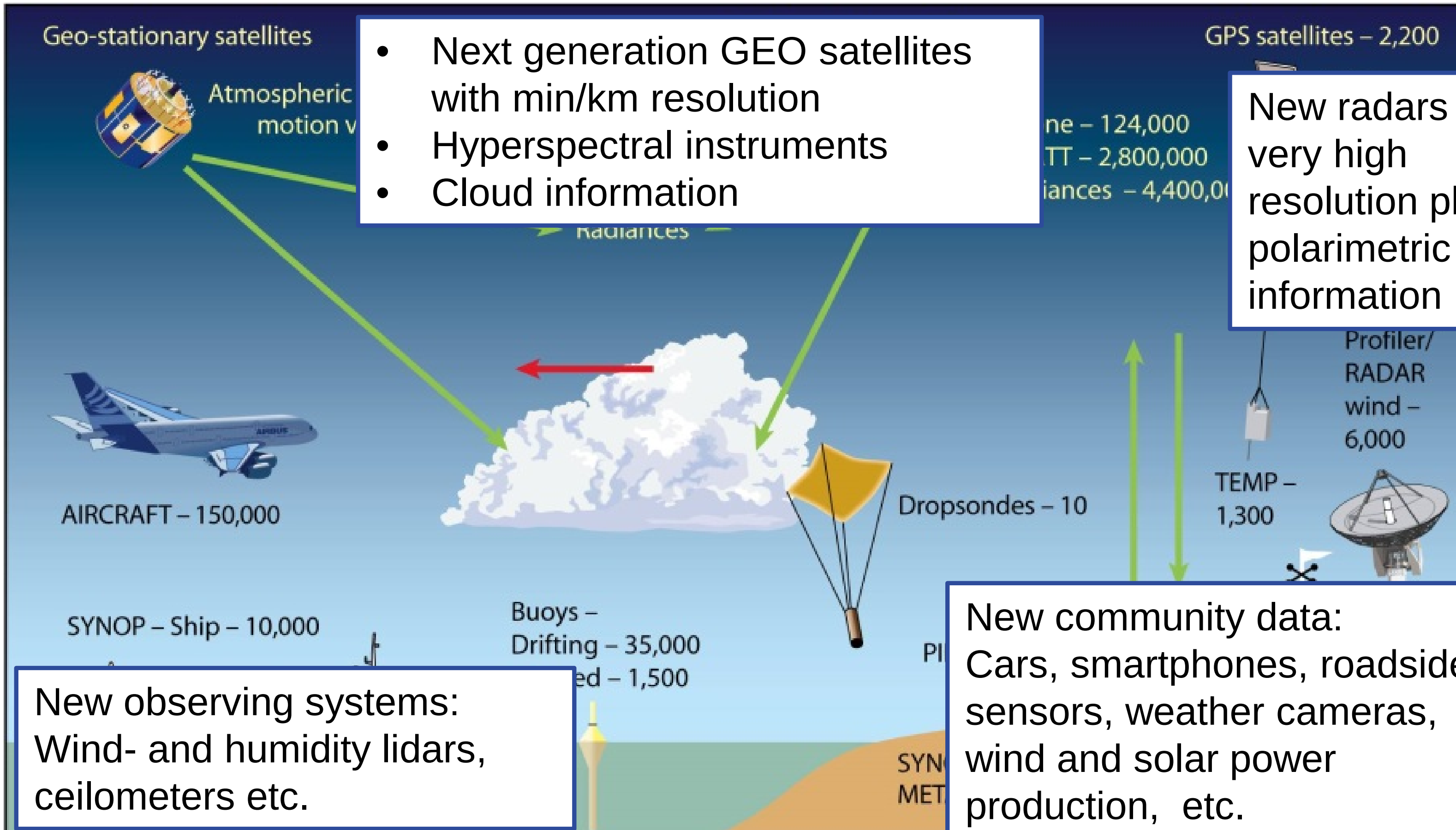
Radar observations

➔ Where should we focus our efforts in data assimilation development?

➔ Which observations do we need in the future?

➔ And how should we approach the two questions above?

Potential observations - challenges for data assimilation



- Next generation GEO satellites with min/km resolution
- Hyperspectral instruments
- Cloud information

New radars with very high resolution plus polarimetric information

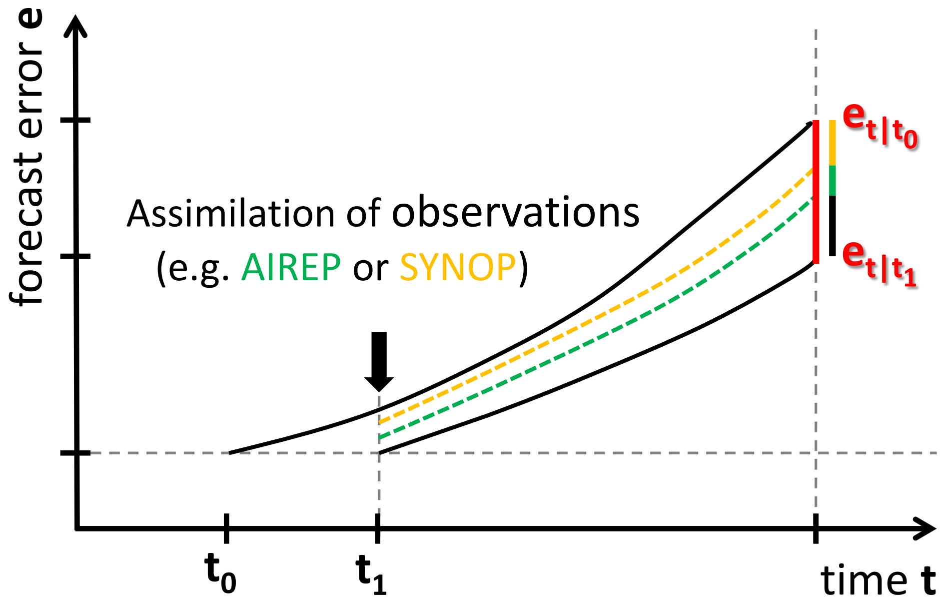
New observing systems: Wind- and humidity lidars, ceilometers etc.

New community data: Cars, smartphones, roadside sensors, weather cameras, wind and solar power production, etc.

→ Huge amount of data → We need to better understand what's need most and at what scale

Part I: The impact of assimilated observation types

Estmating the impact of observations based on ensemble information



Observation Impact ***J(d)***

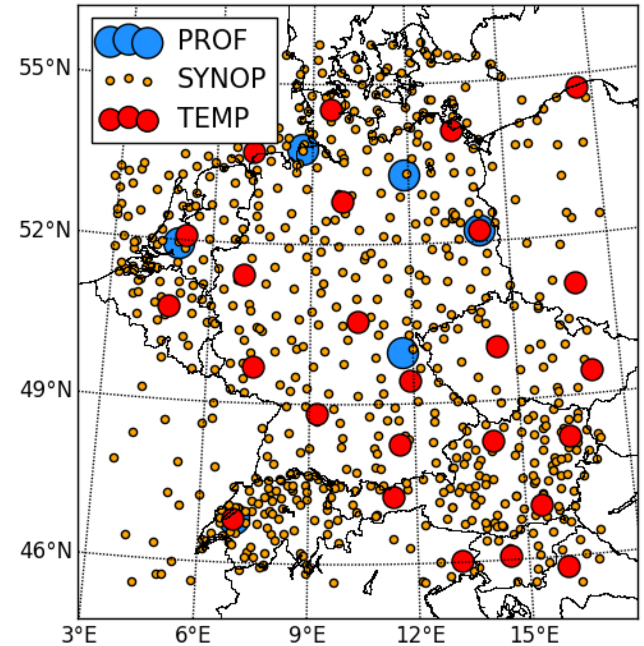
…measured as forecast error difference

$$J(d) \approx \left(\mathbf{e}_f^{\mathbf{d}} + \mathbf{e}_f^{\mathbf{0}}\right)\left(\frac{1}{K-1}X_f^{\mathbf{d}}(X_b W^{\mathbf{d}})^{\mathsf{T}} R^{-1}\mathbf{d}'\right)$$

(Referred to as EFSO, EnFSOI)

Kalnay, E., Y. Ota, T. Miyoshi and J. Liu, 2012: A simpler formulation of forecast sensitivity to observations: application to ensemble Kalman filters. *Tellus A*, 64, 2012.

Experimental setup



Model:

- Regional COSMO-KENDA (LETKF) ensemble system of DWD (2.8 km grid, 40 members)

Setup:

- 3h-cycling
- Verification window: 1-3 h after analysis (longer impact not shown)
- Preoperational setup of DWD without radar LHN

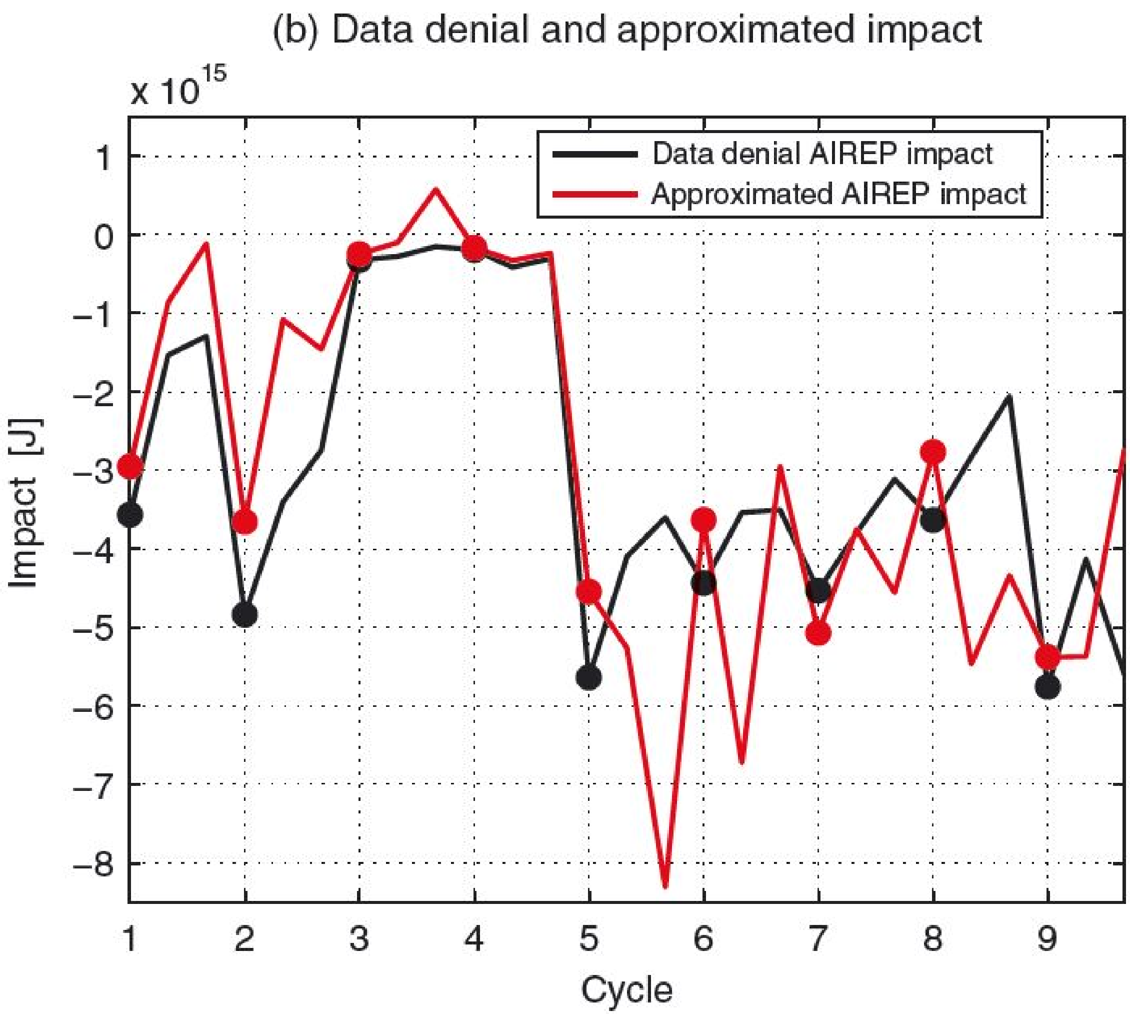
Observations:

- Radiodondes (TEMP), aircraft (AIREP), wind profiler (PROF), surface stations (SYNOP)

Period:

- Three summer periods (2 days, 14 days, 36 days) in 2012, 2014 and 2016 and shorter sensitivity experiments (results not directly comparable)

Validation with data denial experiments



Impact time series of AIREP observations from the **data denial experiment *(black)*** and **approximation *(red)***. Values are displayed for initialization time *(solid circles)* and forecasts up to 6 h *(lines)*.

COSMO-DE / KENDA LETKF
32 member
7 - 9 August 2009

(Sommer & Weissmann 2014)

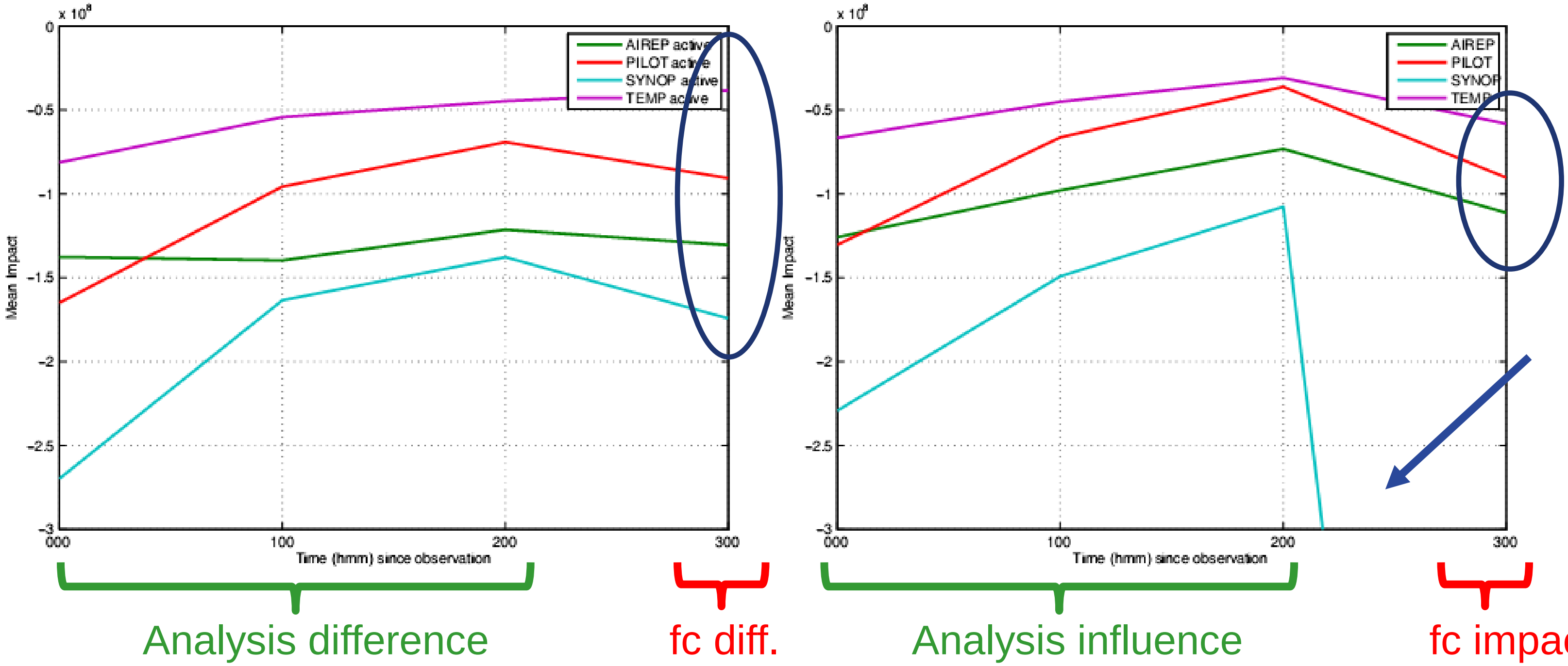
Considerations

Do we now have a plug-and-play tool for calculating observations impact?

No really, because:

- Traditionally, FSO methods use a model analysis to verify a short-range forecast
- The model analysis is highly correlated with the short range forecast
- The calculation is particularly sensitive to biases (as for any verification) and all models have some biases

Observation impact verified with model analyses



Data denial experiment **Estimated impact**

- Reasonable results for AIREP, TEMP and PILOT (i.e. wind profiler)
- Huge impact for SYNOP estimated surface pressure impact at 3h that doesn‘t show up in data denial experiment – related to model bias?
- Different verification as the data denial experiment has no surface pressure observations

Reformulation to verify with observations

$$J(\boldsymbol{d}') \approx \frac{2}{N_e - 1} \boldsymbol{e}_f^d \cdot \boldsymbol{Y}_f^d \left(\boldsymbol{Y}_a^d\right)^T \boldsymbol{R}^{-1} \boldsymbol{d}'$$

J : Observation Impact

N_e: Number of ensemble member

$\boldsymbol{Y}_f^d$: Forecast ensemble in obs . space

$\boldsymbol{Y}_a^d$: Analysis ensemble

$\boldsymbol{R}$: Observation error covariance matrix

$\boldsymbol{d}$: Innovation vector $\boldsymbol{d} = \boldsymbol{y_o} - \boldsymbol{y_b}$

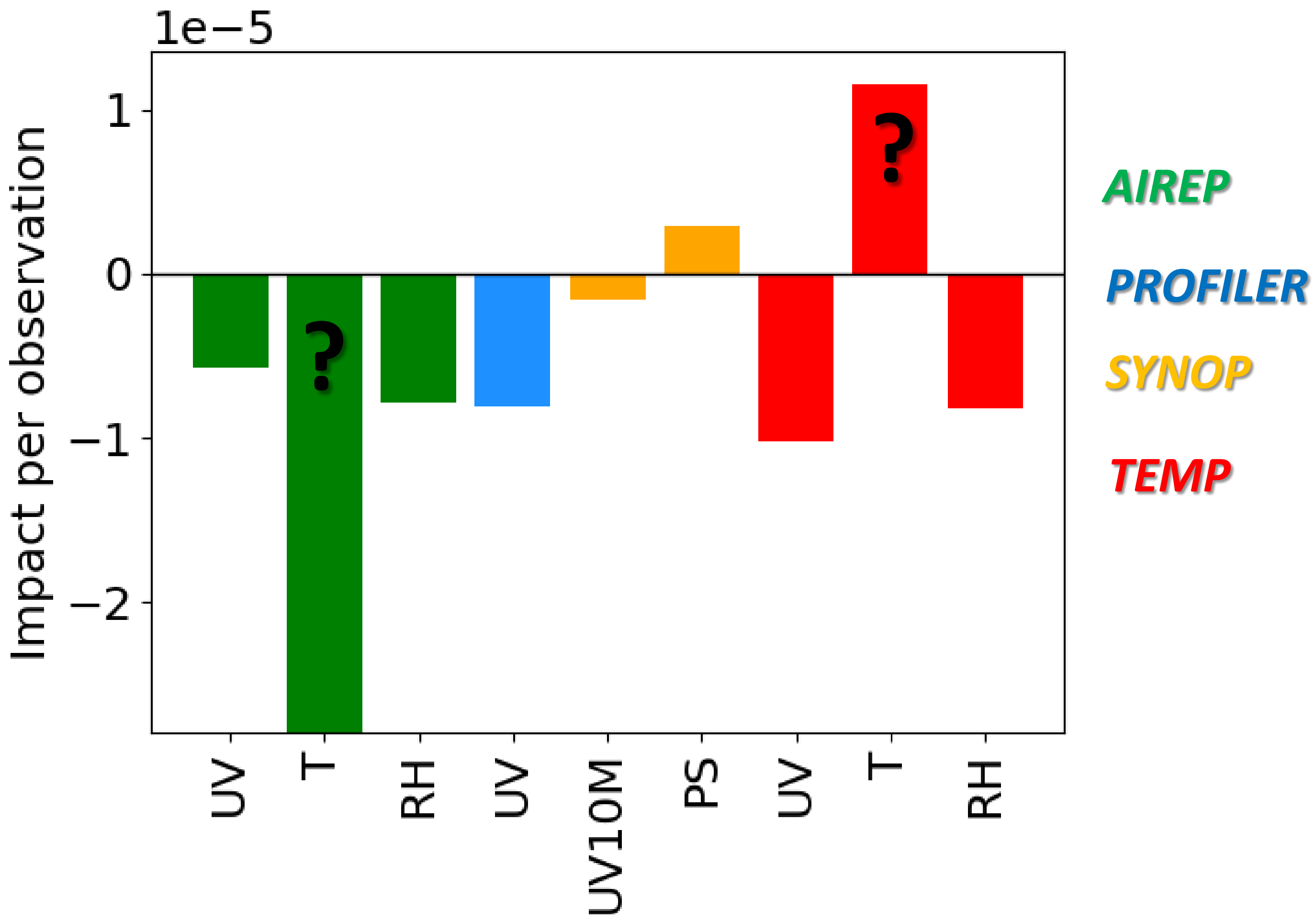
$\boldsymbol{e}_f^d$: Forecast error

Idea:

- Observations should be a better choice for the verification of short-term forecasts

(Sommer & Weissmann 2016)

Impact verified with conv. observations except SYNOP pressure

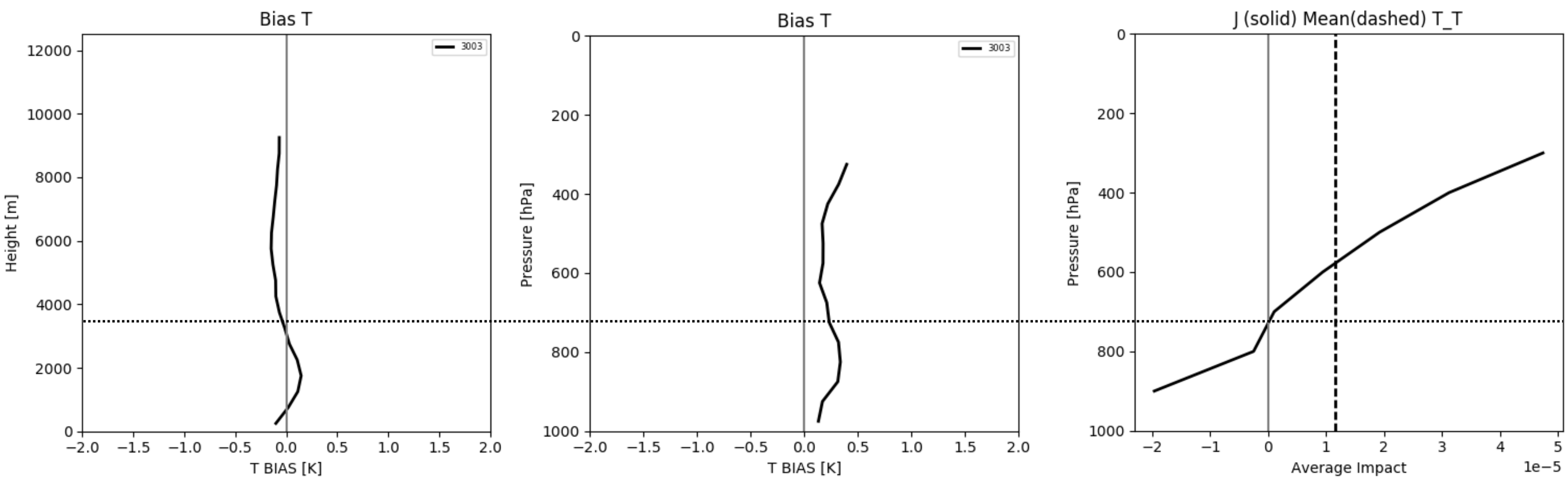


LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN FAKULTÄT FÜR PHYSIK METEOROLOGIE MIM Hans-Ertel-Zentrum für Wetterforschung

Why detrimental impact of radiosonde temperature?

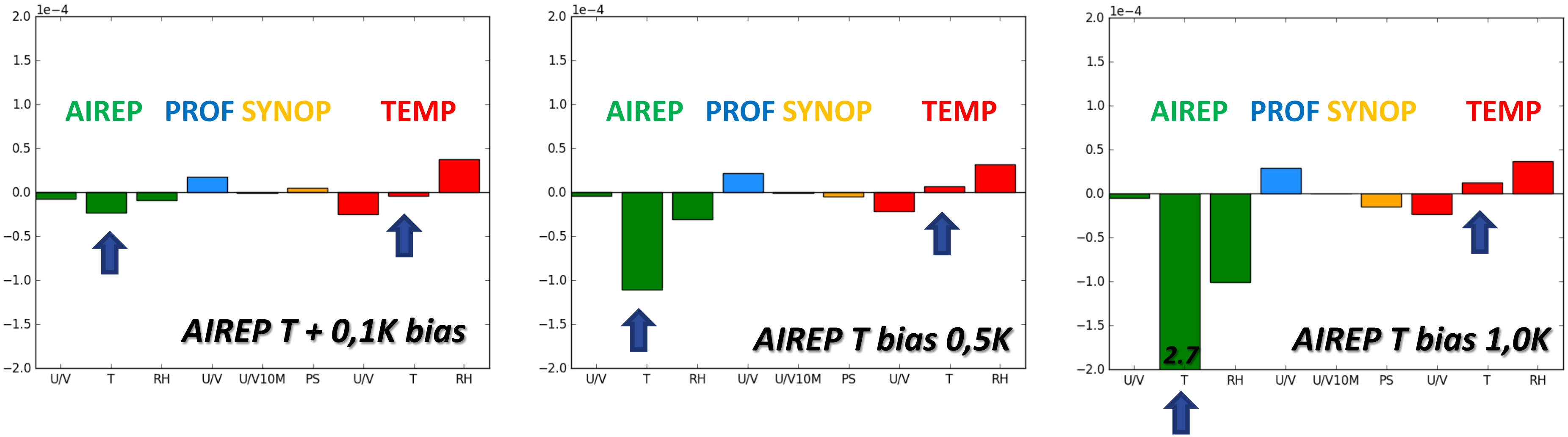
- Radiosondes are usually only launched every 12 h (comparable few observations)
- The forecast impact is therefore mainly verified with aircraft observations

AIREP T bias **TEMP T bias** **TEMP T Impact**



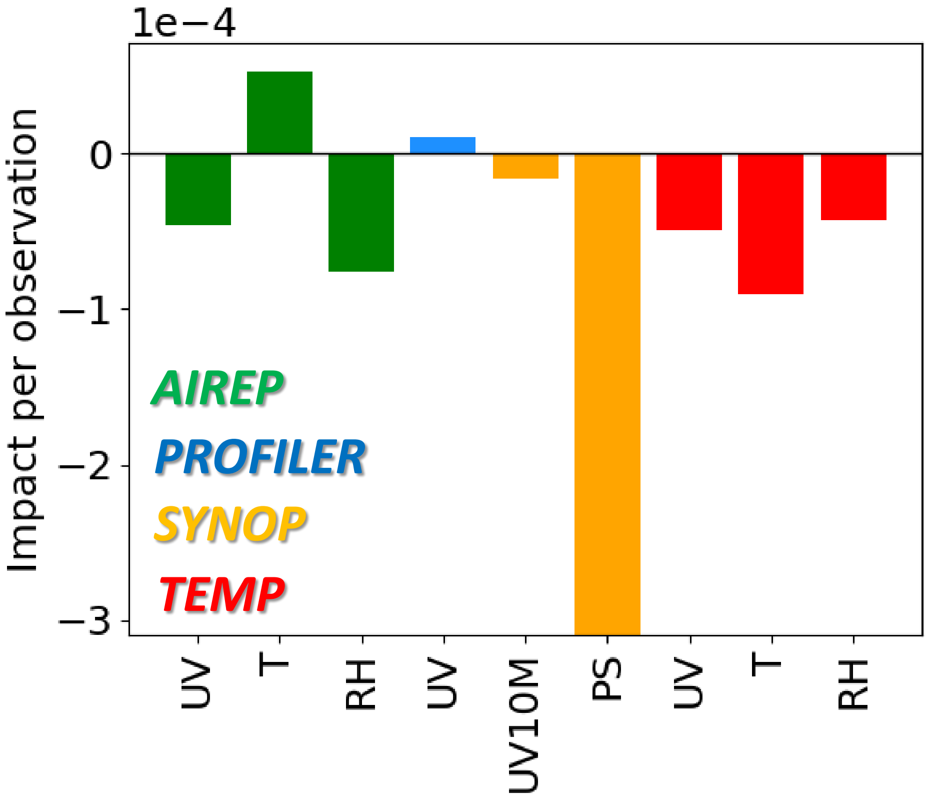
- Presumably radiosondes are unbiased, but the bias correction for aircraft is not optimal
- Even a small/moderate bias significantly influences the results

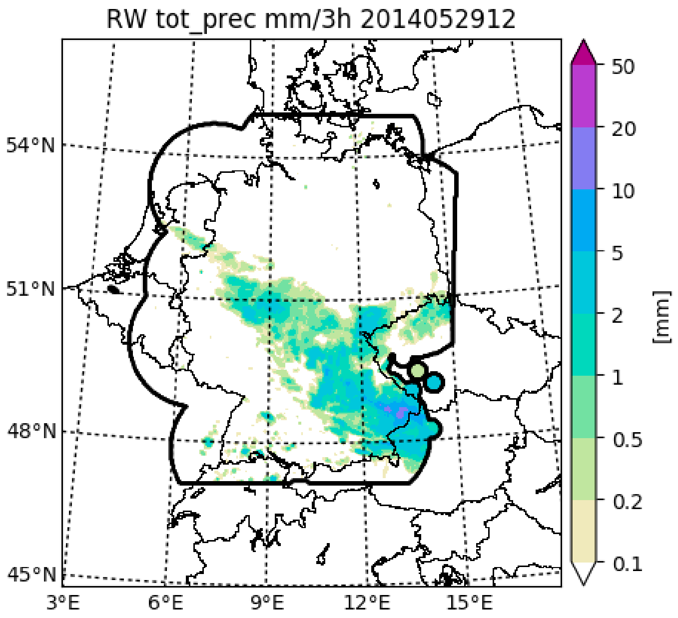
Sensitivity experiment with added bias for AIREP T



- The estimated impact is very sensitive to biases in the verification metric
- Verification should be done with observations independent from the analysis
- Otherwise biases can lead to unrealistic results, especially if there is a correlation of biases between analysis and verification time
- The observed quantity should have a good coverage over the model domain
- Ideally, the quantity used for verification should reflect something that is relevant to the user

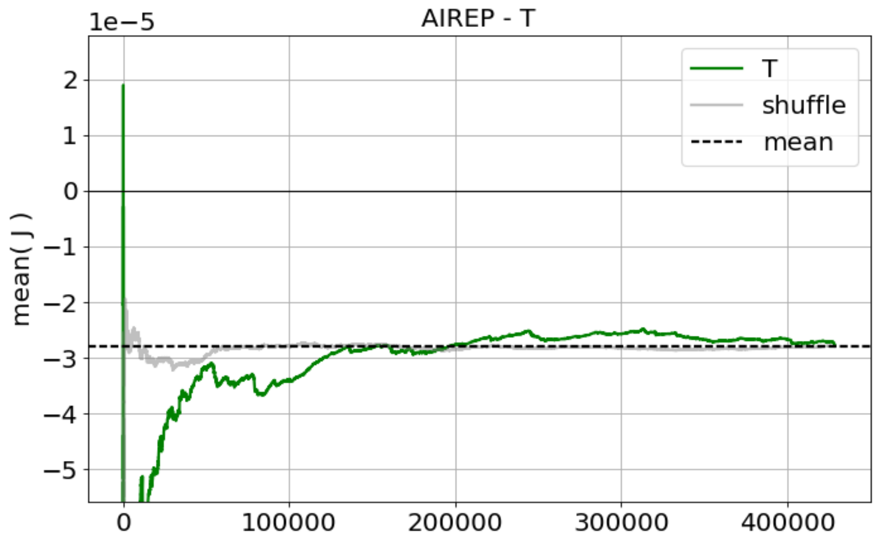
Observation impact verified with radar-derived precipitation

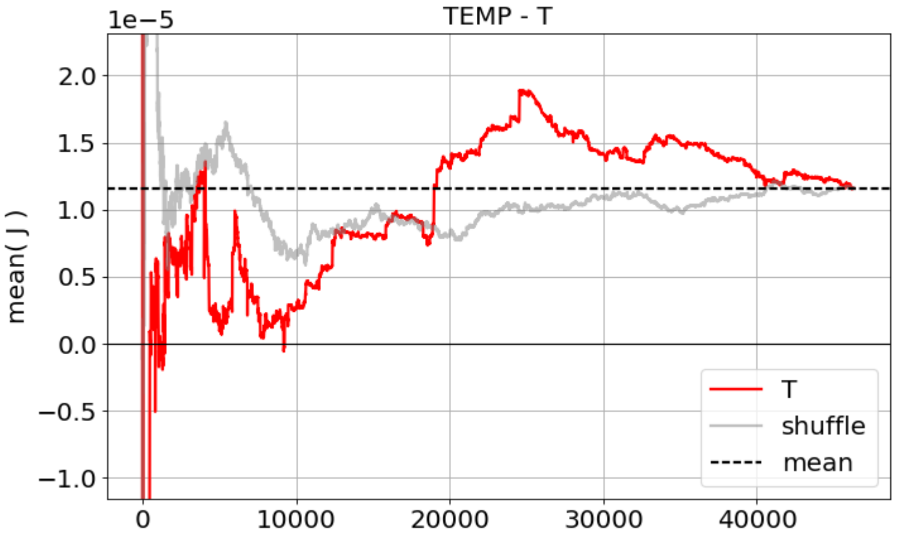




- Large beneficial impact of SYNOP pressure (correction of model bias?), followed by radiosondes and aircraft winds and aircraft humidity
- Detrimental impact of aircraft temperature and wind profiler (related to biases?)

Temporal variability of observation impact verified with conv. obs.





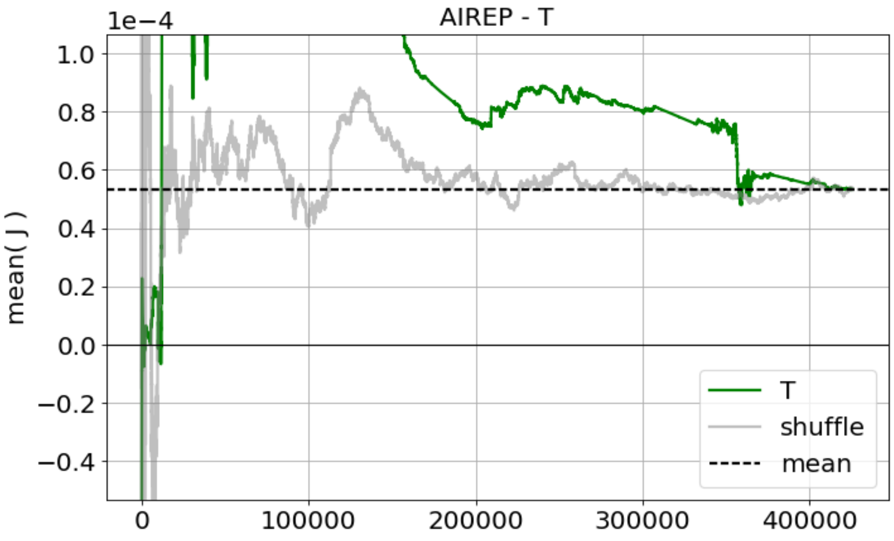
Green/red: Impact averaged starting from the first observation

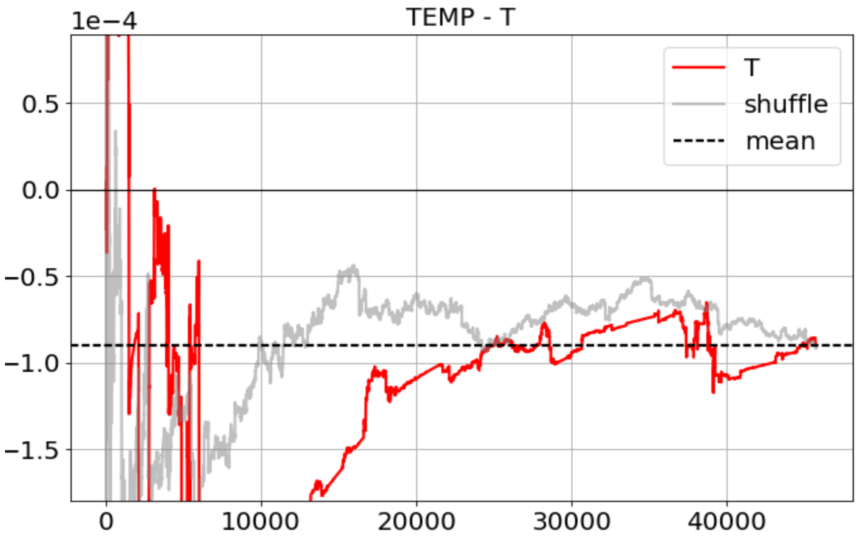
Grey: Impact averaged with random order of the observations

Dashed: Mean impact

Mean impact reached well before the end of the period → representative estimate for this 6-week period, but bias issue mentioned before

Temporal variability of observation impact verified with radar obs.

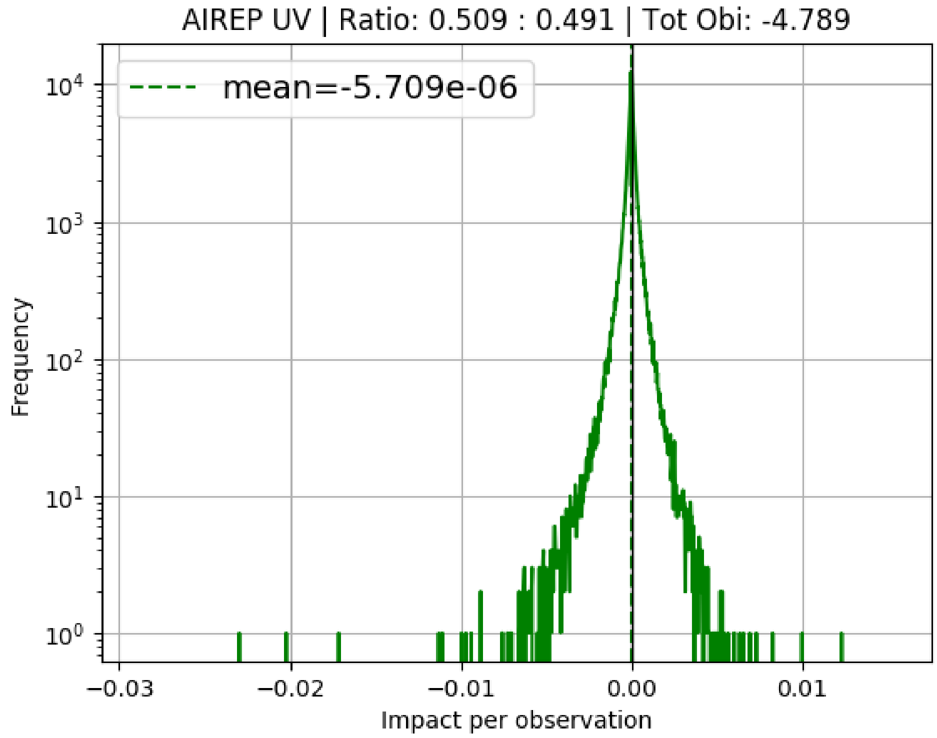




Larger temporal variability of impact verified with radar observations because the amount/area of precipitation varies strongly

Nevertheless, the sign of the mean impacts doesn’t change any more after a few days

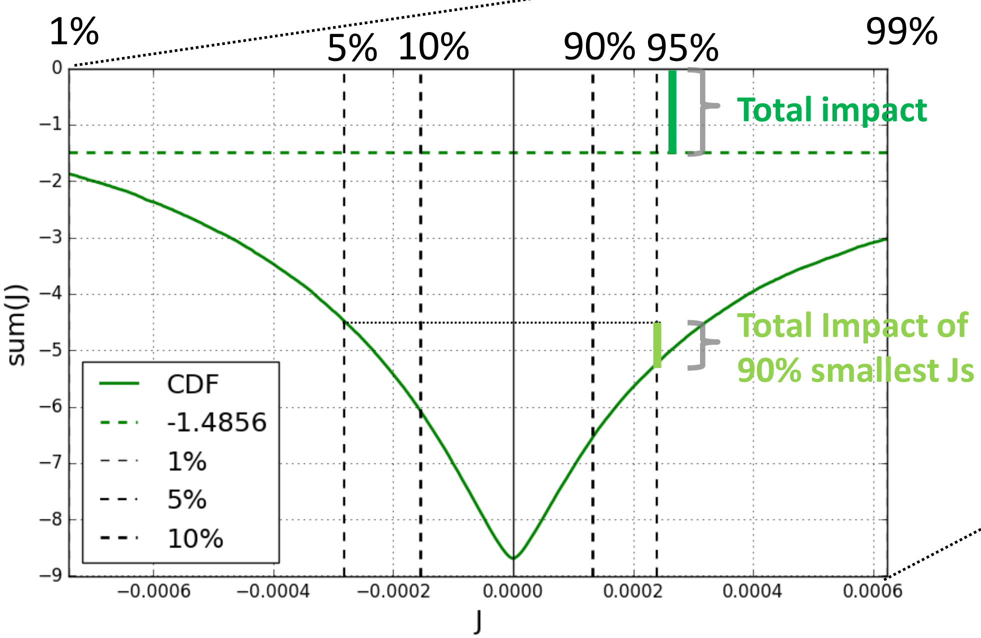
Histogram of observation impact values

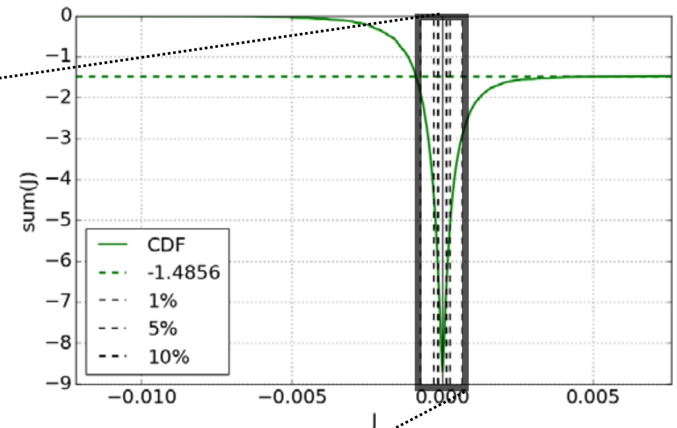


- Observation impact values exhibit a wide distribution and we estimate a small deviation from the mean
- Which observations contribute most to the deviation from the mean? A few large impact observations or many small values?

Question: Is the small number of observations with extreme impact values important?

Cumulative Distribution Function (CDF)



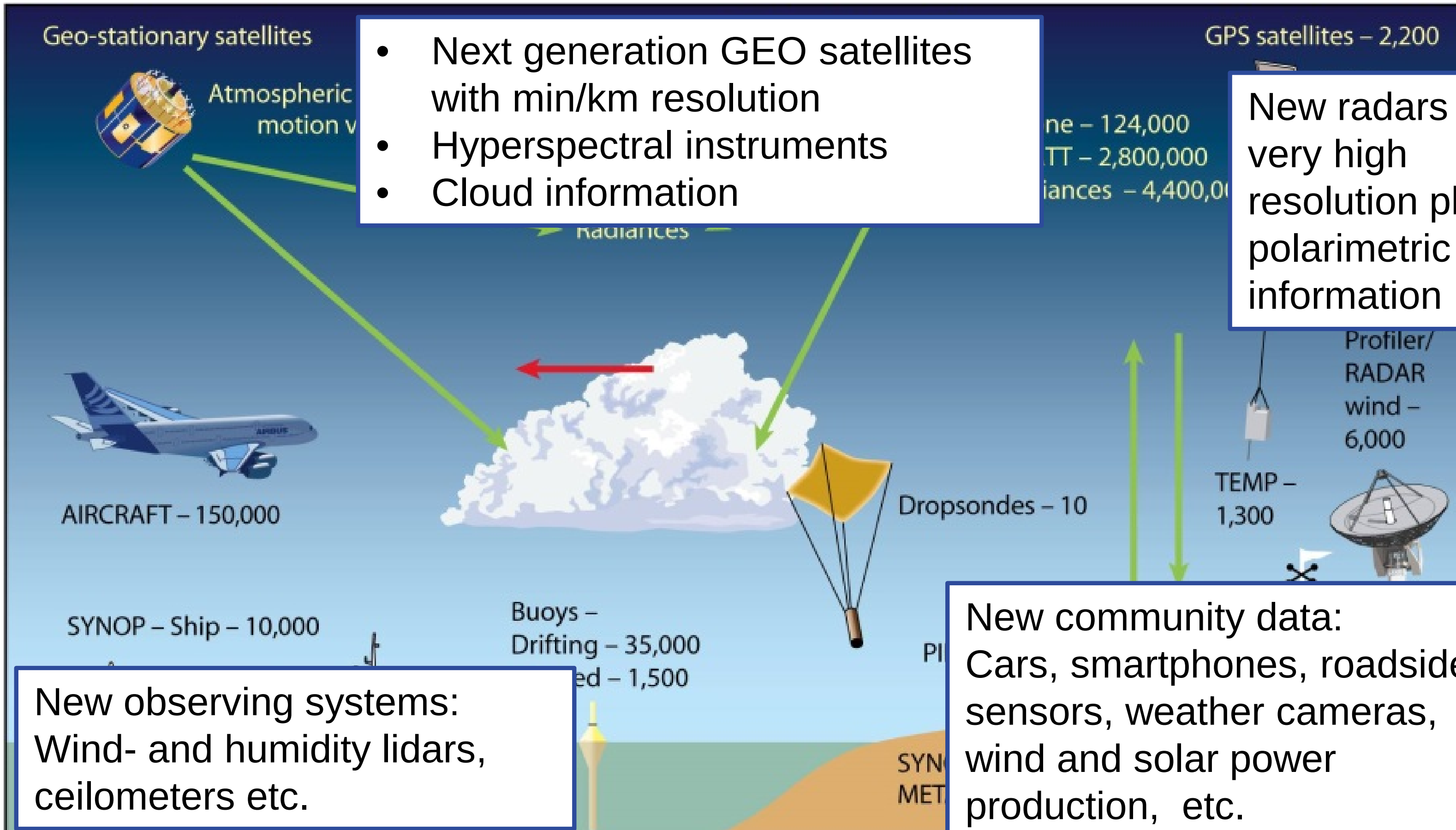


AIREP U/V

Answer: The 10% most extreme impact values contribute as much as smallest 90%

Part II: The potential impact of different observed quantities

Potential observations - challenges for data assimilation



- Next generation GEO satellites with min/km resolution
- Hyperspectral instruments
- Cloud information

New radars with very high resolution plus polarimetric information

New observing systems: Wind- and humidity lidars, ceilometers etc.

New community data: Cars, smartphones, roadside sensors, weather cameras, wind and solar power production, etc.

→ Huge amount of data → We need to better understand what's need most and at what scale

Questions:

- Observations of which quantities are potentially most beneficial for regional DA?
- Which observations (conventional, radar, clouds etc.) are important on which time-scales and spatial scales?

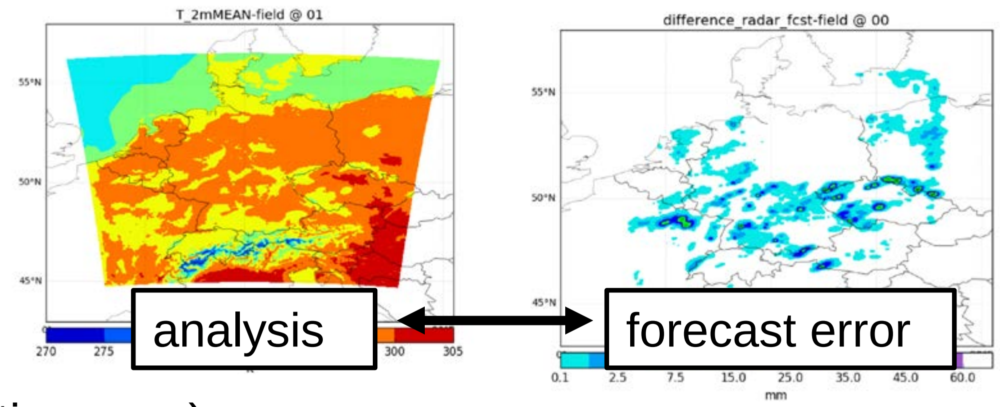
Approach:

- Ensemble Sensitivity Analysis following Torn and Hakim (2008)
- Method has been used in several studies, but mainly to investigate larger-scale dynamics

Ensemble sensitivity gradient

Ensemble sensitivity gradient:

$$\frac{\partial J}{\partial x_i} = \frac{cov(\mathbf{J}, \mathbf{x_i})}{var(\mathbf{x_i})}$$

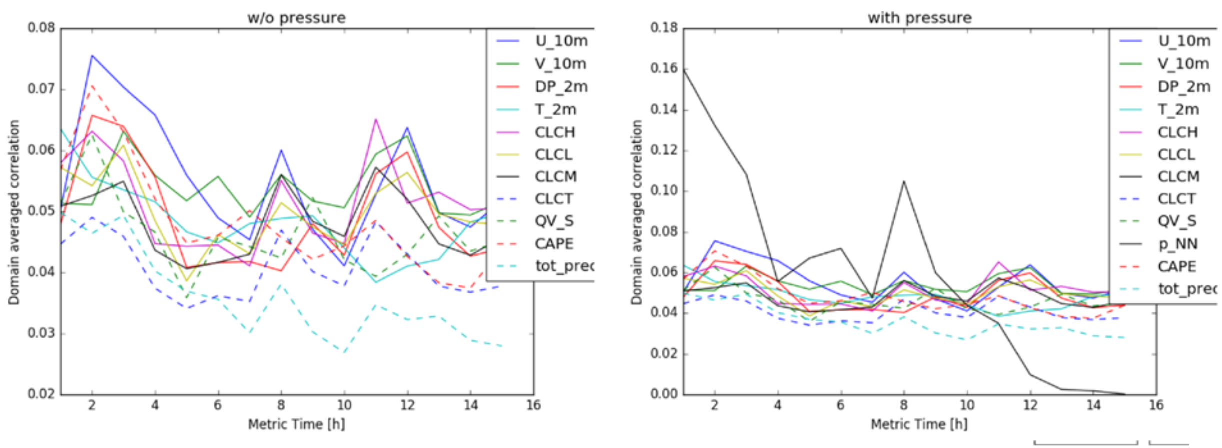


J: Metric of forecast error (e.g. precipitation error)

x_i: Initial condition (T, u/v, q, slp, cloud cover, precipitation etc.)

Following Torn and Hakim (2008) the domain-avaraged forecast sensitivity at each time step can be computed by

$$DAS(t) = \frac{1}{N_h} \sum_{i=1}^{N_h} \left| \frac{\partial J}{\partial x_i} \right| \quad (3)$$



Initial test with a 20-member LETKF ensemble of DWD:

- Large sensitivity for surface pressure, fast decay of sensitivity with time
- However, sensitivities are noisy and spurious correlations may dominate …

→ Cooperation with T. Miyoshi/J. Ruiz to calculate 1000-member ensemble

Conclusions / some thoughts

- Ensemble information can be used to estimate the impact of different observations (known as ensemble FSO/FSOI)
- However, one must be very cautious with potential systematic errors in the verification metric, particularly if this is correlated with the analysis
- We proposed to use observations and ideally independent observations for verification
- As for any other forecast verification, the impact must be evaluated using different metrics to get a complete picture
- We see a large beneficial impact of surface pressure observations – presumably (partly) related to the correction of a model bias. The goal however is to remove the cause of the model bias and not the symptom
- The verification issues (biases, correlation, variable-dependence) have been investigated for the ensemble approach in a convective-scale modelling system, but the same applies for adjoint FSOI and for global modelling systems
- We need better knowledge on where to put our priorities regarding regional observing networks and the inclusion of novel observations in regional data assimilation